1

METHOD AND SYSTEM FOR SENDING AN AUDIO MESSAGE

5

This invention relates to a method for sending an audio message from a sender to a recipient over an audio messaging system and to an appropriate audio messaging system. Further the invention relates to a transmitting device and to a receiving device for such an audio messaging system.

10   The popularity of text-based messaging services has increased immensely since their introduction a few years ago. The widespread Short Messaging Service (SMS) is just one example of such a service. Text news systems like AOL's Instant Messenger, Microsoft's MSM Messenger and Yahoo's Messenger for PCs can be used free of charge after downloading the required free software. Some of these PC-

15   based messaging providers offer a voice-chat functionality in addition to the text messaging services. Furthermore, some other providers have specialised in voice chat, ultimately leading to a voice-over-IP (internet protocol) scenario.

A notable distinction between voice chat functionality and text messaging is the possibility for the user to interact explicitly, for example by choosing a

20   chat window and typing there or by other actions like writing a word document and sending it. On the other hand, voice interaction is continually transmitted, i.e. an uninterrupted interchange takes place. This is often not what the user really wants, for example when he is in a room with other people and only wishes to transmit specific remarks as messages, whereas the remarks directed by him at the other people in the

25   room generally should not be transmitted. Normal telephony allows the user to circumvent this problem by covering the microphone with his hand or switching the telephone to mute. Evidently, this is not possible when using a hands-free telephone or headset. The recipient of a message has a similar problem – while it is possible to read private messages received using text-based messaging services, even when a third party

30   is in the same room, by reading the messages from a screen or a display which cannot be viewed by the third party, it is next to impossible to ensure that audible messages not be heard by third parties for whom the messages are not intended, unless the messages

are listened to through headphones.

Text messaging systems do indeed appear to enjoy a greater acceptance level than voice chat functionality. This is probably owing to the tendency that users do not really desire a permanent conversation experience. On the one hand, they want to be able to connect to the other person. On the other hand, they would just as equally like to be connected in an offline mode in which they are not permanently involved in an on-going conversation in which all their remarks are communicated.

Therefore, an object of the present invention is to provide a method for sending an audio message from a sender to a recipient over an audio messaging system and an appropriate audio messaging system which offers to the user essentially the same experience as text messaging systems. In particular the user should be able to easily send specific utterances as audio messages while excluding other utterances from being sent by the messaging system.

To this end, the present invention provides a method for sending an audio message from a sender to a recipient over an audio messaging system comprising the following steps:

First, a sender's audio message is collected by a transmitting device. The message is usually generated by the sender speaking the message. Nevertheless it is also possible that the sender generate the message or parts of the message in another form, for example by singing, playing on an instrument, clapping hands etc.

This audio message will then be analysed to detect a control information part, also called "audio header" in the following, containing directives such as details of communication specifications of the message; and a main part comprising the effective message or effective information which is to be sent to the recipient, also called "audio body" in the following.

The terms "sender" and "receiver" do not necessarily imply individual users, but can mean user groups, a member or all members of such a group. A user group might use a single shared transmitting or receiving device, for example members of a family to whom the device belongs, or employees in an office using a device designated for that office. A user group might also mean a group of users each of whom

has his own device, in which case a message destined for the user group will be
transmitted to all receiving devices.

The communication specifications of the message, incorporated in the
control information part, may be any kind of transmission and /or presentation

5   specification like, for example, a message type and/or a sending mode, e.g. information
specifying that the message is secret, private, urgent etc. The control information part
could also include information for sender identification or for specifying the recipient
of the message. For instance, a typical audio header may be "Private message from Bob
to Carl". This control information part of the audio message is at least partially

10   interpreted for controlling the audio messaging system for transmitting and/or
presenting the specific audio message. For example, a control signal for the transmitting
device and/or the receiving device and/or other parts like transceiving stations, router
etc. of the audio messaging system may be generated based on the control information
part.

15   In a further step, at least the main part of the audio message is sent to a
receiving device located in the vicinity of the recipient and is presented there to the
recipient.

An appropriate audio messaging system for sending an audio message
from a sender to a recipient according to this method comprises a transmitting device

20   with a user interface for collecting a sender's audio message and message analysing
means for analysing the audio message for detecting a control information part
concerning communication specifications of the audio message and a main part
comprising the actual message which is to be sent to the recipient. Further, the audio
messaging system comprises an interpreting unit for at least partially interpreting the

25   control information part of the audio message for controlling the audio messaging
system for communicating the specific audio message. Additionally, the audio
messaging system comprises a receiving device with a user interface for presenting at
least the main part of the audio message to the recipient. Finally, the audio messaging
system requires a means for transmitting at least the main part of the audio message

30   from the transmitting device to the receiving device.

With the aid of the method and the audio messaging system according to
the present invention, the user controls the audio messaging system by commands

4

embedded in the audio message, thus avoiding a continual transmission of everything he says. In other words, the user can provide the system with "meta-information" in an utterance along with the actual audio content of the message. The system analyses the audio message accordingly and separates the audio header containing the control

5   information from the audio body with any utterances intended for transmission. If the system is unable to detect an audio header with appropriate directions for communicating a message to a particular person in a particular manner, then nothing will be transmitted.

This is illustrated in the following simple example: assuming a user of

10   the system says "Message to Carl: the soccer match starts at 7.00pm", this utterance will be picked up by the user interface of the transmitting device and analysed. The audio header "Message to Carl" will be detected and interpreted, and the message "The soccer match starts at 7.00 pm" will be transmitted to a recipient called "Carl". On the other hand, if the user simply informs another person also present in the room about the

15   start time of the match using the remark "Pete, you know the soccer match starts at 7.00 pm", an activated audio messaging system or the corresponding transmitting device would conclude, on analysis of the utterance, that it does not contain an audio header. The utterance would, as a result, not be identified as an audio message and would not be transmitted.

20   Therefore, the invention provides an exceptionally simple and user-friendly means of controlling the system, so that only certain utterances are transmitted by the audio messaging system to other persons, without having to first deactivate the system or parts of the system, for example a microphone or loudspeaker. Furthermore, the sending user can control the system with respect to transmitting the message and

25   presenting it, whereby all control directives can be comfortably included in the message by means of appropriate formulation in an audio header, without the user having to carry out any manual actions. In other words, the entire control of the audio messaging system can be comfortably carried out using a hands-free set. Thereby, such a system offers advantages over the usual speech control for typical mobile telephones, for

30   example in automotive hands-free sets, whereby a connection to another participant can be initiated and controlled using speech commands, but in which a permanent connection is maintained thereafter between the user and the participant. All of the

5

user's utterances are communicated to the other participant, and muting the telephone is only possible by issuing the appropriate command, or by covering the microphone etc.

The dependent claims and the subsequent description disclose particularly advantageous embodiments and features of the invention.

5   In a preferred embodiment of the invention, the control information part of the audio message is also at least partially transmitted to the receiving device and interpreted for controlling the presentation of the audio message to the recipient. In other words, the receiving device receives appropriate information, with the aid of the audio header, for example as to when, how and to which user(s) the audio message or

10   the audio body of the audio message is to be output. Preferably, the audio header can also be output at least partially to the recipient.

Since the control information part preferably deals with commands spoken by the user, automatic speech recognition techniques can be used to identify the control information part within the audio message, whereby automatic speech

15   recognition in this case does not imply speech recognition in a strict sense, but rather language understanding techniques. To this end, the transmitting device should comprise an automatic speech recognition arrangement.

To assist the identification of the control information part within the audio message, the audio message is preferably built up in a defined composite

20   structure in which the control information part is positioned at a specific position respective to the main part. More preferably, the control information part is positioned at the beginning of the audio message and followed by the main part. The advantage of this is that the control information part is the first to be detected by the speech recognition arrangement, and the following main part need only be buffered or prepared

25   for transmission. The control information part can, however, be located at any suitable position within the message, for example at the end of the message, or the control information part might be distributed over several positions in the message, so that certain control information is located at the start of the message and further control information is located towards the middle or at the end of the message.

30   Analysis of the audio message with the aid of an automatic speech recogniser might involve, for example, searching for certain key-words that might be stored by the audio messaging system in an appropriate memory such as a storage unit

in the transmitting device or receiving device. Typical examples of such key-words might be "message", "message to" etc., descriptors for possible recipients of the messages, as well as key-words specifying the type of message or manner of transmission, for example "secret", "private" or "urgent".

5   To make the transmission of messages as easy as possible, unique identifier strings are associated with the possible users or user groups of the audio messaging system. Such a unique identifier string might comprise, for example, the user's real name, or might equally well be any other string concealing the identity of the various users. In particular, entire user groups can be identified collectively using a

10   single string. The use of nicknames or fantasy names which can most easily be recalled by the other users is preferred. These nicknames are included in the system's vocabulary and can be used to efficiently address a fellow user in the audio header by just saying his nickname. Furthermore, groups can be defined where all connected members will receive the message if the audio header contains the name of the group.

15   Preferably, the identifier strings of the possible recipients are stored together with a corresponding address book entries in a memory of the transmitting device and, if need be, in the receiving device or in a further suitable location in the audio messaging system.

Audio messages will often be sent to a number of people at the same

20   time. During a longer conversation the same list of recipients will be frequently used. When speaking the audio header, it is inconvenient for a user if all names of all recipients have to be spoken each time. Therefore, dynamically associating nicknames or other identifier strings with the list of relevant address book entries will make the sending of messages more comfortable.

25   Preferably a key-word like "Reply" or similar is used to indicate in the audio header that the associated audio message should be transmitted to the sender of the last message received and possibly to all users to whom the last message was sent.

The transmitting device is preferably realised as a dialog system, comprises such a dialog system, or is part of such a dialog system. In this particularly

30   preferred case, an automatic dialog can be initiated between the audio messaging system, or more particularly the transmitting device, and the sender, in order to identify the control information part of the audio message when an ambiguity value (e.g. based

on an internal confidence measure) of a recognition result of the automatic speech recogniser reaches or exceeds a certain ambiguity threshold level.

In other words, if the system is uncertain as to whether a message should be sent, to whom it should be sent, or in which manner it should be sent, the system can issue a prompt to the user asking for confirmation, or can enter into a dialog with the user to allow correction of a supposed audio header. In this way, the system ensures that no message is sent unintentionally, or sent to the wrong recipient.

As already mentioned, the control information part, in a preferred embodiment, is also transmitted at least partially to the receiving device, where it is interpreted to control the output of the audio message. This is particularly useful when information pertaining to identification of the recipient, for example the identifier string, is also transmitted. With the aid of the identifier string, the user can be identified on the part of the receiving device before output of the audio message of the audio body of the audio message takes place.

To this end, in a particularly preferred embodiment, the identifier string of a user or a user group is linked to identifier characteristics of the specific user, user group, or members of a user group. The identifier characteristics can be, for example, a secret sequence of characters, speaker identifier characteristics and/or video characteristics such as the biometric data of the appropriate user. With the aid of these identifier characteristics, the authorised recipient of a certain audio message can be identified from among other possible users present in the vicinity of the receiving device at the time of reception of the message, before outputting the main part of the audio message.

Preferably, the identifier characteristics can be stored in a memory to which the receiving device has access, and the receiving device comprises a means of identifying the recipient on the basis of these identifier characteristics.

One possibility might be that a camera observes the persons present in the room, and identifies the face of the recipient with the aid of the biometric data and using known image processing techniques.

Alternatively, the device might identify the user acoustically. For example, the audio header might be output, followed by an appropriate prompt. If a user answers, he can be identified as the right user by means of speaker identification. The

8

message is only output once the identity of the user has been successfully verified.

In a preferred embodiment, the sender of an audio message can be identified by means of identifier characteristics, and corresponding information regarding the sender can be transmitted along with the audio message. As long as the sender has identified himself in the audio header, for example in the form of "Message from Bob to Carl", it is possible to check the validity of the sender with the aid of the identifier characteristics.

Usually, an audio message should be output immediately to the authorised recipient, on account of topicality. However, there are situations in which the output would be unsuitable, for example when a secret or private message should be output, and the recipient is not alone in the room, or is otherwise occupied and is not able to receive the message. It might be that the recipient is caught up in a conversation or phone-call. Taking account of such situations is particularly important, since an audio message is not enduring. If the user is not in the room or is not paying attention, and the message is output immediately, it would be irretrievably lost.

To this end, a preferred method according to the invention automatically analyses the situation in which an identified recipient is currently involved, and the audio message is presented to the recipient in a specific form and/or at a specific time depending on the situation. For example, if the recipient is present and not engaged in an absorbing task (such as a telephone conversation), an incoming message can be played immediately. Otherwise the message can be buffered and played as soon as the user enters the room or concludes his task. If an interruption of longer messages is necessary (e.g. due to an incoming phone-call) playback can be resumed at a later point in time.

There are different methods of automatically analysing the situation in which the recipient is currently involved. In a preferred embodiment, a very satisfactory receiving device is realised as a dialog system with the additional ability to receive pictures of its environment by means of a camera or similar device. The identity of the recipient and/or the current situation could then be determined by using known image processing techniques. A very easy method of identifying the recipient and/or analysing the current situation is to initiate an automatic dialog between the audio messaging system/receiving device and the recipient. For example the device could precede the

9

dialog described above by outputting the audio header "Message for Carl", and then issuing the prompt "Are you ready to receive the message?". Should the user reply with "Yes", the message will be presented, otherwise it will be buffered until the user explicitly requests the message at a later time.

As already described above, the audio messaging system, besides the transmitting device located in the vicinity of the sender, also requires a receiving device located in the vicinity of the actual recipient.

A suitable transmitting device should comprise at least the following components:

- a user interface for collecting a sender's audio message;

- message analysing means for analysing the audio message for detecting a control information part concerning communication specifications of the audio message, and a main part comprising the effective message which is to be sent to a specific recipient;

- an interpreting unit for at least partially interpreting the control information part of the audio message which controls the audio messaging system with respect to communicating of the audio message;

- a transmitting interface for transmitting at least the main part of the audio message to a receiving device.

A suitable receiving device should comprise at least the following components:

- a receiving interface for receiving an audio message sent by a transmitting device and comprising a control information part concerning communication specifications of the audio message and a main part comprising the effective message sent to a specific recipient;

- a user interface for presenting at least the main part of the audio message to the recipient;

- an interpreting unit for at least partially interpreting the control information part of the audio message which controls the audio messaging system with respect to presentation of the audio message.

As already explained above, the transmitting device and/or the receiving device are preferably realised as dialog systems. The transmitting device and receiving device can be constructed identically and can comprise all necessary components for

10

transmitting as well as receiving messages. Dialog systems used for other purposes such as control of other devices can be equipped with appropriate components, so that such a dialog system can be used as transmitting device and/or receiving device for an audio messaging system according to the present invention.

5   In an especially preferred embodiment, the transmitting device and the receiving device comprise part of a dialog system such as that described in DE 102 49 060 A1. In this case, the dialog system need only be further equipped with an appropriate message analysing means, an interpreting unit and a transmitter/receiver interface in order to be able to transfer audio messages via a communication network.

10   The message analysing means might be essentially the speech recognition unit already present in this device, supplied with the appropriate vocabulary for detection of the audio header. An interpreting unit for interpreting the control information part of the audio message can preferably be realised as a software routine within the actual dialog control unit, or in a different form of software running on a processor of the dialog

15   system. The interpreting unit must be able to convert the control directives contained in the audio header into control signals, so that the message is sent in the intended manner from the sender's transmitting device to the receiving device of the recipient, or that the received message is presented in the correct manner to the right recipient by the receiving device.

20   Other objects and features of the present invention will become apparent from the following detailed descriptions considered in conjunction with the accompanying drawings. It is to be understood, however, that the drawings are designed solely for the purposes of illustration and not as a definition of the limits of the invention.

25

Fig. 1 is a schematic diagram showing one embodiment of an audio messaging system according to the invention;

Fig. 2 is a perspective view of a preferred embodiment of the

30   transmitting and/or receiving device for an audio messaging system according to Figure 1;

Fig. 3 shows a very simple example for an audio message with a

11

structure according to the invention;

Fig. 4 is a flow chart which shows a process flow in a transmitting
device commencing with user input up to transmission of the audio
message.

5

Figure 1 shows a audio messaging system with, for the sake of
simplicity, only two devices, namely a transmitting device 2_T in the vicinity of the
sender U_S, and a receiving device 2_R in the vicinity of a recipient U_R, where the
10   transmitting device 2_T and the receiving device 2_R are connected to each other by means
of a network N.

The communication network N can be any kind of network, such as a
telephone network, a mobile telephony network, the internet, an office intranet or a
home-communication network. It is only necessary that the two devices 2_T and 2_R can
15   communicate with each other by means of appropriate interfaces 14.

Generally, such an audio messaging system 1 comprises a considerably
greater number of devices. Any number of devices might be incorporated. In particular,
it is not necessary that a certain message only be sent from one particular device to
another device. Such a message can be sent simultaneously to several devices, for
20   instance to send a message from one user to a user group, i.e. to many recipients.

In the example shown, the transmitting device 2_T and the receiving
device 2_R are generally constructed in the same manner, i.e. they can be used for both
receiving and transmitting audio messages. The references 2_T and 2_R only serve to
distinguish between receiving device 2_R and transmitting device 2_T for the sake of
25   clarity. In general, a message can also be transmitted in the opposite direction.
Therefore, to simplify matters, the devices will also be referred to as "transceiving
devices" 2_T, 2_R where appropriate.

Such a transceiving device 2_T, 2_R is constructed in an advantageous
arrangement as a dialog system.
30   A dialog system of this kind comprises, along with other components not
shown in the figure, a user interface 10 with an arrangement for picking up or collecting
audio signals from a user such as speech or singing, by means of a microphone or

12

something similar. This user interface 10 also features an acoustic output arrangement 12, such as a loudspeaker. Furthermore, the user interface 10 can comprise components for visual output or input, such as a display and/or a camera.

In a preferred embodiment, shown in Fig. 2, the user interface is

5   moveable, for example can rotate about an axis, and mounted on a housing 18, which might contain any further components of the transceiving device 2_T, 2_R. The user interface 10 has a clearly recognisable front aspect 17, comprising a loudspeaker 12, two microphones 11, and a camera 16. Furthermore, this embodiment might comprise a display unit (not shown in the figure) for visual output of information. A preferred

10   dialog system with such a display unit is the home dialog system described in DE 102 49 060 A1, which is incorporated herewith in its entirety. The additional functionality advantageous for the present invention and achieved with such a realisation of the transceiving device 2_T, 2_R, is explained at a later point.

Further components of the transceiving device 2_T, 2_R are an audio

15   control unit 8, which, for example, controls the audio functions of the user interface 10 and prepares incoming speech signals for later processing steps. An example of such a later processing step is an automatic speech recognition arrangement 7, comprising an actual speech recognition unit 5 followed by a subsequent language understanding unit 6. With the aid of these components, the incoming speech signals of the user U_S can be

20   analysed and recognised in the usual manner, i.e. the underlying meaning of the spoken input can be determined.

The speech recognition results are then forwarded to the dialog control unit 3, which controls the actual dialog with the user, and works together with an application – in this case a message transceiving application 12 – in order to send or

25   receive an audio message. This message transceiving application 13, along with a physical network interface 14 connecting to the communication network N, ensures that the message can be sent and received in an appropriate electronic form. The message transceiving application 13 together with the network interface 14 can therefore also be regarded as a "receiving interface" or "transmitting interface" or also as a "transceiving

30   interface" as appropriate.

Since output to the user is necessary to allow a dialog with the user U_S, U_R, the system also features a prompt generator 9 for generating output prompts. Such a

13

prompt generator 9 can output pre-generated prompts retrieved from a memory, or can comprise a speech generation unit for converting text prompts into speech signals, which can be output as synthetic speech by means of the audio controller 8 and the user interface 12.

5   An audio message of a sending user U_S can be sent to a recipient U_R, in this case another individual user, in the following manner:

The sender U_S speaks the audio message AM which is detected by the user interface 10, or more precisely the audio detection arrangement 11, of the transceiving device 2_T. The recorded speech signals are then pre-processed by the

10   audio control unit 8 and forwarded to the kernel of the automatic speech recognition unit 5, which analyses the utterance of the user U_S together with the subsequent language understanding unit 6.

According to the invention, such an audio message AM comprises a control information part CP (audio header) along with the actual information to be

15   transmitted which is the so-called main part MP. This structure is shown in Figure 3. The message shown here "Private message to Carl: the meeting starts at 7.00pm" contains the control information part CP "Private message to Carl", followed by the main part MP "The meeting starts at 7.00pm".

The automatic speech recognition arrangement 7 is configured in such a

20   way that it can identify the control information part CP and separate this from the main part MP. To this end, the vocabulary of the automatic speech recognition arrangement 7 contains certain control words CW, which, if they occur within a certain syntax, will be identified as belonging to a control information part CP of an audio message AM.

These control words CW are stored in a memory unit 15 within the

25   receiving device 2_T. Furthermore, this memory unit 15 also stores identifier strings IS, such as nicknames of various users of the audio messaging system which might be possible recipients. A corresponding "buddy list", containing nicknames of potential recipients and their addresses within the audio messaging system 1, can be assembled by the user of the transmitting device 2_T. This list can be stored in the transmitting

30   device 2_T or at another location of the audio messaging system 1, for example on a server of a service provider.

In the example shown in the figures, both main part MP and control

14

information part CP of the audio message AM are passed from the automatic speech recognition arrangement 7 to the dialog control module 3, in which an interpreting unit 4, for example in the form of software routines, is installed. This interpreting unit 4 also has access to the control words CW and identifier strings IS in the memory 15, and

5   therefore can interpret the control information part CP of the audio message AM in order to generate corresponding control signals for the audio messaging system 1, particularly the transmitting device 2 $_T$, and thus to control the audio messaging system 1, particularly the transmitting device 2 $_T$, accordingly. If the control information part CP is not clearly identifiable, the dialog control unit 3 initiates a dialog by, for example,

10   causing the prompt generator 9 to issue an appropriate prompt to the sender U$_S$, for instance "Are you trying to send a private message to Carl?". The sender U$_S$ can answer with a simple "Yes" or "No", as appropriate, either to confirm a presumed control header CP, or to terminate the procedure in the case of an erroneously detected control header CP.

15   If the system has ascertained that a control header has been correctly identified, or if the user has confirmed a presumed control header through an ensuing dialog, the main part MP of the audio message AM, attached to the audio header CP, is sent to the recipient U$_R$ specified in the audio header CP by means of the identifier string IS, which in the case of the preceding example is the user with the nickname

20   "Carl".

            To this end, the dialog control unit 3 passes the main part MP, and preferably also the control information part CP, to the message transceiving application 13 and simultaneously passes on any corresponding control signals, so that the audio message AM can be communicated, via the communication network N, to the address

25   of the receiving device 2$_R$ of the user with the nickname "Carl". The control information part CP and the main part MP of the audio message AM are then transmitted to the receiving device 2$_R$ via the network interface 14 connected to the communication network N.

            The sequence of operation within the transmitting device 2$_T$ is shown in

30   the flow chart of Figure 4. The process commences at step I with the user input. In step II, an appropriate analysis determines whether the user input comprises an audio header CP, whereby the following step III checks to see if all the required parts of an audio

15

header are present and clearly identifiable. Otherwise, step IV initiates a dialog, i.e.
questions are put to the user and the answers are analysed until all the required parts of
an audio header have been identified. A typical case of misinterpretation might arise
with the following: "Private message to Julie: Ann, shall we meet for lunch today?".

5   This message might be interpreted to give the audio header "Private message to Julie"
and the main part "Ann, shall we meet for lunch today?" or the audio header "Private
message to Julian" and "Shall we meet for lunch today?". In this case the system may
prompt "Did you want to send a private message to Julian?" The sender U_S can reply
"No, I wanted to send a private message to Julie". Here, the answer clarifies the

10   misinterpretation by specifying the first of the possible alternatives. In step V, the audio
body, i.e. the main part MP, can be separated from the audio header CP. Subsequently,
further processing steps are possible within the dialog. In the example above, the user is
asked whether further information is the be sent with the audio message AM, i.e.
whether an image or a video is to be transmitted. Other attachments might equally

15   accompany the audio message AM, such as a document. If the user confirms, the
processing step VII can determine which image or video is to be attached to the
message. Another prompt in step VI can ask whether any more pictures, videos etc. are
to be attached. Once the message is complete, step VIII concludes transmission of the
message.

20   At the receiving device 2_R, the control information part CP and the main
part MP of the audio message AM are received over the network interface 14 and
processed by the message transceiving application 13 in the device. Output of the
message is performed by the dialog control unit 3, if necessary the prompt generator 9,
and the audio control unit 8 as well as the loudspeaker 12 of the user interface 10 of the

25   receiving device 2_R.

         To avoid output of the message if the intended recipient U_R is not in the
room, is otherwise occupied at the time, or is in the company of other persons for whom
the contents of the message are not intended, the receiving device 2_R analyses the
situation in advance. For example, the moveable user interface (see Fig. 2) might swivel

30   about in order to scan the entire room with the aid of the camera 16. Using known
image processing techniques, it can be determined whether the intended recipient U_R is
present in the room. The intended recipient U_R can be identified with the aid of the

16

identifier characteristics IC associated with the various identifier strings IS stored in the memory.

To this end, the identifier string IS accompanying the message are used by the message transceiving application 12, or a similarly suitable module of the receiving device 2_R, to retrieve the corresponding identifier characteristics IC from the memory 15 and to identify the recipient U_R using these identifier characteristics IC. The identifier characteristics IC might be biometric data used in image processing to identify the recipient U_R from among other persons in the room.

Equally, speaker identifications characteristics can be applied. In this case for example, the dialog control unit 3 can ensure that only the audio header CT — "Private message to Carl" — is output via the audio control unit 8 and the user interface 10 of the receiving device 2_R, followed by the supplement "Would you like to listen to the message right away?", generated by the prompt generator 9. When the user thus addressed replies, the spoken answer can be analysed in turn by the speech recognition unit 5 and the language understanding unit, and simultaneously checked for validity by speaker identification, whereby extracted characteristics are compared with the information characteristics IC in the memory 15, to determine whether the right user and authorised recipient U_R is answering.

Furthermore, with the aid of the camera 16 and usual image processing techniques, it can be determined whether the user is involved in a conversation with other users, whether he is making a phone call, or is involved in any other situation making him unable to receive the message.

If the recipient U_R is not in the room, or not able to receive the message AM, the message is buffered and output at a later point in time. If the recipient U_R indicates that he would like to listen to the message in privacy, the receiving device 2_T will also buffer the audio message AM and not play it until the recipient U_R is alone again in the room, or until the recipient U_R has ensured that he will able to privately listen to the audio message AM, for example by wearing headphones or similar.

The user interface 10 of the receiving device 2_R advantageously turns to present its front aspect 17 to the authorised recipient of the message, recognised by receiving device 2_R, i.e. the receiving device 2_R turns to directly face the recipient U_R when outputting a dialog prompt or the audio message AM or the main part of the audio

message AM. Other advantageous means of outputting or usage of the receiving device 2_R or transceiving device 2_T, realised in the form of a dialog system, are described in the document DE 102 49 060 A1.

Although the present invention has been disclosed in the form of
5   preferred embodiments and variations thereon, it will be understood that numerous additional modifications and variations could be made thereto without departing from the scope of the invention. In particular, the transmitting device and/or the receiving device might, for example, be constructed using a different architecture than that described.

10   For the sake of clarity, it is also to be understood that the use of "a" or "an" throughout this application does not exclude a plurality, and "comprising" does not exclude other steps or elements. A "unit" may comprise a number of blocks or devices, unless explicitly described as a single entity.